

MUsa un modelo de evaluación de Entornos virtuales de enseñanza y aprendizaje. Aplicación a un caso de estudio.
[pág. 94-103]

MUsa un modelo de evaluación de Entornos virtuales de enseñanza y aprendizaje. Aplicación a un caso de estudio

Ariel Ferreira Szpiniak¹, Cecilia V. Sanz²

¹ Departamento de Computación. Universidad Nacional de Río Cuarto, Río Cuarto, Córdoba, Argentina

² III LIDI, Facultad de Informática - Universidad Nacional de La Plata, La Plata, Buenos Aires, Argentina

Resumen

En este trabajo se presentan los últimos avances en la construcción de MUsa, un modelo de evaluación de Entornos Virtuales de Enseñanza y Aprendizaje (EVEA) centrado en la usabilidad. MUsa se compone de cuatro capas que parten de una evaluación general (capa 1) hasta la evaluación más específica y cercana al usuario final (capa 4), de un EVEA. Las cuatro capas persiguen objetivos diferentes. La primera es la encargada de realizar una evaluación del entorno en general, las tres capas restantes se sitúan en un contexto de uso particular, por ejemplo el aula virtual de un curso. Para ello, las capas más cercanas al usuario proponen escenarios de uso que permiten guiar y sistematizar la evaluación. Se presentan en este artículo las ideas centrales en las que se basa MUsa. Se describen cada una de las capas que lo conforman y aspectos vinculados a su aplicación. Finalmente, se detallan algunos resultados y líneas de trabajos futuros.

Palabras clave: EVEA, usabilidad, evaluación.

Abstract

The latest advances in the construction of MUsa, a model for assessing Teaching and Learning Virtual Environments (TLVE), centered on usability are presented. MUsa consists of four layers that go from a general assessment (layer 1) to a more specific assessment that is closer to the end user (layer 4), of a TLVE. The four layers have different purposes. The first layer is responsible for the general assessment of the environment, while the other three layers are located in specific contexts of use, such as a virtual

classroom. To do so, the layers that are closer to the user propose use scenarios that allow guiding the assessment using some sort of system. In this article, the main ideas behind MUsa are presented. Each of the layers is described, as well as aspects related to their application. Finally, some results are detailed and future lines of work are mentioned.

Keywords: TLVE, usability, assessment

1. Introducción

Los Entornos Virtuales de Enseñanza y Aprendizaje (EVEA) forman parte del conjunto de aplicaciones informáticas diseñadas para la utilización de Internet con fines educativos. Su principal característica es la interactividad, como estrategia para favorecer el contacto entre docentes, alumnos, y materiales de aprendizaje. En términos generales, suelen ser versátiles para poder adecuarse a diferentes propuestas y procurar que el diseño tecnológico acompañe al modelo pedagógico. En realidad, los EVEA coadyuvan a la concreción de objetivos educativos al proveer una serie de herramientas que facilitan la gestión de usuarios y cursos, y los procesos de comunicación, evaluación, colaboración, y distribución de contenidos [1]. Presentan una serie de funcionalidades para lograr que los procesos de enseñanza y aprendizaje, puedan desenvolverse de forma mediada acorde a las necesidades de cada contexto específico. Sin embargo, para poder determinar en qué medida se logra ese objetivo es necesario realizar una evaluación.

La evaluación puede orientarse de diferentes maneras, según lo que se pretenda evaluar. En general, los modelos existentes analizan los EVEA desde el punto

de vista funcional, sin tener en cuenta la forma en que se puede poner en práctica todo el potencial del EVEA. Se trata de modelos donde el usuario final no participa, sino que son expertos quienes realizan la evaluación, asumiendo los diferentes roles para los cuales el entorno será utilizado. Esto se debe a que dichos instrumentos fueron pensados principalmente para la toma de decisiones sobre el entorno a adoptar. Por ello, en la mayoría de los casos, se realiza una extensa categorización de características deseables o ideales que debería tener un sistema de este tipo. En la actualidad, los EVEA que han logrado buenos niveles de aceptación, disponen de gran parte de las características buscadas en estos modelos de evaluación. Es de destacar que en dos de los siete modelos estudiados está presente la usabilidad como uno de los ejes de evaluación: el propuesto por Adell [2], que lo explicita claramente, pero lo acota en el análisis, y el de Pere Marquès [3], que trabaja sólo con algunos aspectos relacionados con la interfaz. Por tal motivo, se cree necesario incorporar una forma de evaluar, donde no se pierdan de vista estas características, pero se prioricen otros aspectos centrados en el modo en que los usuarios finales toman contacto con cada funcionalidad. En este sentido, la usabilidad es un concepto muy útil, ya que tiene estrecha relación con la utilidad funcional. La usabilidad busca analizar cuán bueno es un sistema como para satisfacer todas las necesidades y requerimientos de los usuarios [4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18]

Las evaluaciones de usabilidad son factibles de utilizar en diferentes tipos de sistemas y distintas fases del proceso de desarrollo de software. Muchos de los métodos o técnicas existentes pueden aplicarse en más de una fase, aunque, la mayoría está destinada a capturar problemas de usabilidad en etapas tempranas de desarrollo. Otra manera, es aquella que se ocupa de evaluarla como un atributo del producto final. Posiblemente, el hecho de centrarse en el contexto de uso sea uno de los rasgos más fuertes de esta mirada, ya que pone el acento en la utilización concreta que el usuario hace del sistema, analizando en ese contexto sus características. El contexto de uso involucra a: los tipos de usuarios, las tareas, el equipamiento, y el entorno físico y social, dentro del cual será utilizado el sistema. Para poder evaluarlo es necesario entonces contar con usuarios reales, situaciones reales de trabajo y condiciones específicas. Por lo tanto, el producto deberá estar concluido o en una etapa avanzada de desarrollo. El usuario y el entorno en el cual se utilizará el sistema pasan a ser ejes centrales para evaluar la usabilidad del mismo.

El modelo de evaluación de usabilidad que se propone, denominado MUsa, está orientado a una evaluación de producto, es decir, sobre EVEA que ya han sido desarrollados, y están en pleno funcionamiento. La evaluación se apoya en escenarios reales de uso, teniendo especial consideración por los alumnos y docentes, los objetivos que estos se proponen, las tareas específicas que realizan dentro del entorno durante las actividades de enseñanza y aprendizaje, el modelo mental que utilizan, el equipamiento e infraestructura del que disponen, el lugar físico donde habitualmente se desenvuelven, y el entorno social en el cual están insertos. MUsa también puede facilitar la evaluación de nuevas versiones de un mismo sistema o la comparación entre ellas.

2. Descripción general de MUsa

MUsa propone analizar las herramientas y posibilidades funcionales del EVEA (objeto de evaluación), pero además, indagar sobre su usabilidad desde la mirada de expertos, trabajando con el sistema concretamente, y la de los usuarios reales en los contextos de uso propios para los que fue diseñado. Es por ello, que establece una estrategia basada en cuatro niveles o capas de evaluación, que parten de lo general para llegar a lo particular, es decir, una estrategia top-down [19]. La primera capa es la encargada de realizar una evaluación del entorno en general, desde una análisis más técnico, y las tres capas restantes se sitúan en un contexto de uso particular, involucrando diferentes usuarios, que van desde aquellos que poseen una mirada experta en este tipo de sistemas a los usuarios finales con diferentes tipos de características y contextos de uso (ver figura 1).

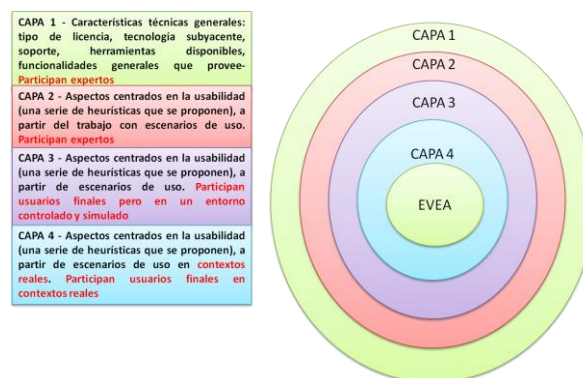


Figura 1. Gráfico representativo de MUsa con las 4 capas y las características principales de cada una

Así, las capas más cercanas al usuario utilizan escenarios de uso que permiten guiar y sistematizar la

evaluación. Para cada capa, se han planteado métodos y técnicas de recolección de datos diferentes, que se complementan entre sí, y posibilitan enriquecer la evaluación. Las capas propuestas pueden ajustarse a distintos contextos dependiendo de la necesidad, nivel de profundidad de la evaluación, recursos y tiempo disponible.

A continuación se describe la intencionalidad de cada una de dichas capas.

2.1 Capas que conforman el modelo MUa

La idea de conformar el modelo de evaluación en capas surge de la necesidad de abordar diferentes niveles de análisis, que parten de aspectos que se vinculan con una mirada más superficial del objeto de estudio, a otra más centrada en sus posibilidades concretas para los usuarios finales de éste, y los contextos para los que fue creado. Para ello, se han planteado cuatro capas que permitirían idealmente conformar una evaluación completa de este tipo de aplicaciones.

2.1.1 Capa 1

Está destinada a analizar aspectos relacionados con la aceptabilidad práctica del EVEA. La aceptabilidad práctica conjuga la utilidad (usefulness) con otras variables como costos, compatibilidad, confiabilidad, y soporte, entre las más importantes [20]. La utilidad a su vez está determinada por la utilidad práctica o funcional (utility), la usabilidad, y accesibilidad. En la primera capa, se aborda el análisis de todos los componentes de la aceptabilidad práctica, menos el referido específicamente a usabilidad. Se analizan cuestiones vinculadas a accesibilidad (parte de la utilidad), junto con las características técnicas generales (antecedentes, potencial, tecnología utilizada, licencia, soporte, seguridad, acceso de usuarios), compatibilidad (servidor, usuario/cliente, formatos multimedia, integración, seguimiento de estándares) y robustez (integridad de funcionamiento, recuperación ante fallos, seguridad). Así, se pone énfasis sobre las características funcionales del EVEA, donde se consideran cuestiones relacionadas con las facilidades para la organización académica y flexibilidad pedagógica [19]. Un punto que toma especial relevancia es la versatilidad del entorno para adaptarse al desarrollo de cursos, grupos de cursos, carreras, comunidades virtuales, etc. También es importante analizar la forma en que puede dar soporte a diferentes modalidades educativas, sean estas a distancia o mixtas, y a diferentes enfoques de enseñanza y de aprendizaje. Para ello, es necesario contar con especialistas en educación y en tecnología. Esta primera capa reúne varios de los elementos que se proponen en diferentes métodos de evaluación de

EVEA, revisadas a partir del estudio bibliográfico de la temática [19, 21].

El instrumento de evaluación diseñado para esta capa, se focaliza en los componentes tecnológicos de los EVEA, pero sin perder de vista los componentes organizativos, pedagógicos y didácticos, ya que ambos son determinantes al momento de analizar un entorno particular [22].

2.2 Capa 2

Está orientada a evaluar la forma en que el sistema interactúa con el usuario, la interfaz que presenta, y el modo en que permite a los usuarios realizar las tareas básicas. Se proponen los métodos de inspección, como el recorrido cognitivo y la evaluación heurística, como los adecuados para analizar estos elementos [23, 24, 25, 26, 27, 28, 29, 30]. Para ello se han definido para esta capa una serie de heurísticas de usabilidad organizadas en dos niveles de evaluación: uno general, orientado a la tarea, y otro particular, orientado al diseño. El primero examina el aspecto y comportamiento desde el punto de vista de las tareas y objetivos de los usuarios, y el segundo, los aspectos concretos del conjunto de interfaces provistas por el EVEA para completar las tareas.

En esta segunda capa de MUa se requiere de la participación de expertos, y como se mencionó, se utilizan métodos de inspección basados en heurísticas y recorridos cognitivos para evaluar la forma en que el sistema interactúa con el usuario, la interfaz que presenta, y el modo en que permite realizar las tareas básicas [31]. Ambos métodos resultan similares en cuanto a la forma en que se trabaja con expertos, escenarios, y preguntas que debe hacerse el evaluador durante la aplicación, es por ello que se han conjugado sus posibilidades, de manera de abordar un análisis más completo. Así, el experto transitará una serie de escenarios de uso adoptando distintos personajes, navegando por las interfaces respectivas a los fines de determinar el grado en que son respetadas las heurísticas, tanto desde el punto de vista de los propósitos como desde los objetivos de los usuarios de un EVEA. Para MUa se han definido los escenarios de manera tal de permitir describir una situación de uso real, en la cual el experto pueda tomar el lugar del usuario final. Se construyeron 21 escenarios a partir de un contexto de uso determinado, donde los personajes principales compuestos por un docente responsable, dos docentes tutores, y tres alumnos, realizan las tareas más importantes y frecuentes relacionadas con dicho contexto [31]. La definición de tareas se fundamentó a partir de una serie de encuestas realizadas con los usuarios de este tipo de sistemas, en la que se rescataron las tareas más frecuentes que éstos abordan.

2.3 Capa 3

Las dos primeras capas focalizan su atención en la mirada que los expertos pueden hacer sobre el EVEA. En esta tercera capa el objetivo es nutrirse de la opinión del usuario final, pero dentro de un ambiente controlado, y con la participación de observadores que guíen y faciliten el proceso. Para ello, es imprescindible la participación de usuarios reales. La forma elegida para llevar a cabo esta tercera etapa de evaluación es mediante la instrumentación de test de usuarios [32]. El objetivo es establecer en qué medida el EVEA se adapta a los estilos de trabajo reales de los usuarios, en lugar de forzar a los usuarios a adaptar sus estilos de trabajo. Dentro de las alternativas existentes para este tipo de métodos, se ha seleccionado una variante del test de pensamiento en voz alta, denominada test de expresión del usuario en base a preguntas [24]. La metodología consiste en brindarle al usuario un escenario tipo, y solicitarle que efectúe las tareas involucradas en dicho escenario, bajo la atenta mirada de un observador. A medida que el usuario interactúa con el EVEA, el observador debe realizar preguntas directas acerca del producto o la tarea que el usuario está realizando, mientras que éste debe expresar en voz alta sus pensamientos, sensaciones y opiniones. El observador debe tomar registro de todo ello para procesarlo una vez concluido el test. Para ello, se distinguen dos momentos de trabajo en esta capa: evaluación directa e indirecta, la primera es realizada por los usuarios, y la segunda por los observadores. Para guiar a los usuarios en el tipo de información a revisar, se han planteado una serie de heurísticas de usabilidad. Estas serán descritas en detalle en la sección 3 de este trabajo. Aquí se han elegido específicamente, aquellas orientadas a la tarea, tales como complejidad, visibilidad, intuitividad, y topografía natural. También se incorporan las métricas relacionadas con efectividad, eficiencia, y satisfacción. Los usuarios participan de la evaluación realizando una o más tareas que forman parte de un escenario tipo. Es importante que los usuarios involucrados en los test abarquen los diferentes roles en que puede interactuarse con el EVEA. Por otro lado, dentro de cada rol deben seleccionarse usuarios con diferente nivel de experiencia en el uso del servicio Web. Esta capa es muy costosa en términos de tiempo, cantidad de participantes y análisis, puesto que requiere de observaciones directas de distinto tipo de usuarios trabajando sobre el entorno. Si bien no existe consenso sobre cuál es la cantidad óptima de evaluadores a incorporar, es necesario buscar un equilibrio para que sea factible de realizar, 10 como

mínimo y 30 como máximo acorde a lo aportado por algunos autores [7]. En tal sentido, para obtener un abanico de usuarios que cubra gran parte del espectro deben seleccionarse 10 usuarios, como mínimo, de los cuales 4 deben ser docentes y 6 alumnos, con el objetivo de cubrir 2 roles de tutor y los 3 de alumno, con al menos dos usuarios diferentes de cada uno. En la evaluación directa, el usuario final debe situarse en cada uno de los escenarios de uso diseñados e interactuar con el entorno realizando tareas como si estuviera participando de un curso o asignatura mediada por el entorno virtual. Para ello, utiliza un guión provisto por el observador. Luego de concluida la sesión, debe completar un cuestionario. Para la evaluación indirecta, el observador debe estar presente durante toda la sesión y registrar lo que acontece mientras el usuario realiza las tareas. El observador debe realizar preguntas directas al usuario, y éste debe expresar en voz alta sus observaciones, sugerencias, u opiniones. El usuario también puede realizarle preguntas al observador. La evaluación indirecta pretende recabar información sobre un conjunto de métricas de usabilidad que intentan medir el grado de efectividad, eficiencia y satisfacción del EVEA. Para ello el observador debe tomar nota del rol y escenarios que posee cada usuario, comentarios, sugerencias u opiniones del mismo, y las métricas. Se han diseñado para MUso todos los instrumentos de recogida de datos pertinentes a esta capa del modelo.

2.4 Capa 4

Está destinada exclusivamente a que los usuarios finales aporten su punto de vista dentro de un contexto o ambiente real. Para realizar este tipo de evaluación también son adecuados los test de usuarios. En particular, los test remotos son muy eficaces, rápidos y fáciles de realizar [33]. Además, el usuario realiza el test en su propio medio o ambiente, con lo cual es posible evaluar el contexto de uso. Los test remotos se basan, principalmente, en el uso de cuestionarios para recolectar la información. Las heurísticas coinciden con las utilizadas en la segunda capa para la evaluación orientada a la tarea (complejidad, visibilidad, intuitividad, y topografía natural), y se agregan otras vinculadas al diseño en general (productividad, retroalimentación, control por parte del usuario, reversibilidad y manejo del error, diseño y organización, consistencia, ayuda y documentación, y estándares). El evaluador está encargado de suministrar el cuestionario a los usuarios reales, examinar las aulas virtuales luego de finalizadas las tareas, y posteriormente procesar los resultados. También completa un cuestionario que

contempla las métricas de efectividad, eficiencia y satisfacción.

Recapitulando, las dos primeras capas utilizan los expertos para evaluar el EVEA, mientras que la tercera y la cuarta hacen lo propio, pero con usuarios reales. En particular, la tercera capa involucra a los usuarios a partir de un ambiente controlado, con la participación de observadores que guían, facilitan y registran el proceso. Esta cuarta, y última capa, hace lo propio pero dentro de un contexto real de uso, sin laboratorios de usabilidad, ni observadores, es decir, en un ambiente totalmente natural donde intervienen otras variables como la ubicación geográfica, puntos de acceso, tipos de equipamiento y acceso a la red, tipos de software y versiones, además de las tenidas en cuenta en la tercera capa.

Al igual que en las capas anteriores es importante que los usuarios involucrados en los test abarquen los diferentes roles y posean distintos niveles de experiencia en el uso del servicio Web.

Para brindar flexibilidad al modelo de evaluación es posible aplicar la última capa de dos maneras diferentes. La más sencilla es hacerlo sobre un EVEA que se esté utilizando, es decir, que se base en el uso real del entorno virtual, por parte de usuarios alumnos y docentes, en uno o varios cursos que tengan cierta semejanza con los escenarios de uso definidos anteriormente. La otra alternativa es recrear en el EVEA los escenarios que hemos definido, es decir, generar espacios virtuales para contener cursos e invitar a participar a docentes y alumnos. En este caso, el curso no es real sino un simulacro, donde el contenido de aprendizaje no tiene relevancia, pero los participantes son alumnos y docentes que trabajarán en sus propios contextos. La evaluación real es una alternativa simple para cuando los recursos humanos son escasos, se necesita una evaluación más rápida. Este tipo de evaluación presenta una serie de desventajas porque en cursos reales pueden no estar presentes algunas tareas. Además, la cantidad de evaluadores no es la misma que para una evaluación simulada, y tampoco existe garantía de que estén representados todos los tipos de usuarios. Para reducir estos riesgos deben seleccionarse cuidadosamente los cursos y asegurarse que la cantidad de usuarios sea representativa, tanto en cantidad como en experticia.

2.5 Ajuste de MUsa a diferentes contextos

Cada capa del modelo plantea el acercamiento hacia el usuario de acuerdo al objetivo que persigue, siendo la primera la más lejana y la cuarta la más cercana. Las capas pueden ajustarse a distintos contextos dependiendo de la necesidad, nivel de profundidad de la evaluación, recursos y tiempo disponible. Puede prescindirse de alguna, o utilizarlas todas, además de

seleccionar diferentes alternativas dentro de la segunda, tercera y cuarta capa. Las alternativas planteadas con el fin de poder aplicar el MUsa en diferentes situaciones y contextos son:

- Capa 1 → Capa 2 → Capa 3 → Capa 4 (completo)
- Capa 1 → Capa 2 → Capa 4
- Capa 2 → Capa 3 → Capa 4
- Capa 2 → Capa 4

Las versiones reducidas del modelo, constituyen alternativas con el fin de reducir costos y tiempo, pero se considera que el recorrido por cada una de las cuatro capas es la opción más adecuada para una evaluación completa del EVEA.

3. Heurísticas y métricas de usabilidad planteadas en MUsa

Los instrumentos utilizados en la segunda, tercera, y cuarta capa del modelo presentan una serie de heurísticas y sub-heurísticas de usabilidad que guían tanto a expertos como a usuarios a aportar la información requerida para la evaluación del EVEA. Las heurísticas están organizadas en dos niveles de evaluación, uno general, orientado a la tarea, y otro particular, orientado al diseño [34], tal como se ha mencionado en la capa 4. Todas las heurísticas han sido seleccionadas a partir de una revisión bibliográfica específica [21]. La segunda y la cuarta capa utilizan todas, mientras que en la tercera solo participan las orientadas a la tarea.

- **Evaluación de alto nivel orientada a la tarea:** examina el aspecto y comportamiento desde el punto de vista de las tareas y objetivos propuestos por los escenarios de uso. Está compuesta por 4 heurísticas y 17 sub-heurísticas. Las heurísticas son:
 - **Complejidad:** Dificultad u obstáculos encontrados.
 - **Visibilidad:** Diseño natural. Información de contexto y orientación. Acciones posibles en cada momento. Información justa y necesaria para cada tarea. Jerarquización. Señales que indiquen de manera natural lo que se debe hacer. Ayuda de forma implícita.
 - **Intuitividad:** propuesta de un buen modelo conceptual, que permita reducir la necesidad de recordar. Incentivar el reconocimiento. Posibilitar búsquedas.
 - **Topografía natural:** Metáforas. Analogías con el mundo real. Familiaridad. Claridad. Objetos autoexplicativos.

- **Evaluación en detalle orientada al diseño:** analiza los aspectos concretos del conjunto de interfaces provistas por el EVEA para completar las tareas. Contempla una mayor cantidad de heurísticas, 8 en total, junto a 88 sub-heurísticas. Las heurísticas son:

- **Productividad:** Facilidad de realización de tareas habituales. Reducción del número de interacciones a usuarios frecuentes o expertos. Herramientas para trabajo masivo. Recordatorios o avisos de cambios producidos.
- **Retroalimentación:** Informar el resultado de las acciones. Informar los resultados exitosos en forma diferente a los fracasos. Respuesta modesta ante acciones frecuentes y de poca importancia y elocuente ante acciones infrecuentes e importantes.
- **Control por parte del usuario:** Participación e interacción con el sistema. Autonomía. Predicción de los efectos de los actos realizados. Control del sistema en todo momento. Libertad para actuar. Reconocimiento de sucesos. Salidas distinguibles. Prevención de errores.
- **Reversibilidad y manejo del error:** Cancelación, revocación de acciones. Prevención de errores. Detección de errores. Reconocimiento del error. Mensajes visibles y claros de error. Mecanismos de recuperación y solución ante el error. Puntos de salida. Contacto con asistencia.
- **Diseño y organización:** Diseño natural. Edición con posibilidad de ver el aspecto real. Anticipación a las necesidades. Legibilidad. Interfaz visible. Mostrar trabajo pendiente. Acciones relacionadas agrupadas. Página principal como sumario. Información relevante únicamente. Reducción del número de imágenes. Información importante en la parte superior. Uso de lenguaje simple. Párrafos breves. Enlaces visibles. Información importante resaltada. Información relacionada agrupada. Longitud de página acotada. Ausencia de enlaces rotos. Mapa del sitio. Fácil exploración. Diseño general uniforme de estructuras y colores. Diseño atractivo. Ausencia de distractores. Colores armónicos.
- **Consistencia:** Disposición precisa y jerárquica. Título significativo en navegador. Enlace que refleja el título. Similares términos para similares ítems. Hipervínculos claros. Iconos entendibles. Secuencias constantes de acciones se repiten en situaciones similares. Misma terminología en

avisos, menús, y pantallas de ayuda. Comandos constantes se emplean en todas partes del mismo modo. Normas y convenciones usuales de la plataforma Web en cuanto a íconos, botones, imágenes, ventanas y diálogos.

- **Ayuda y documentación:** Documentación suficiente. Ayuda en línea y orientada a las acciones del usuario. Facilidad de búsqueda. Ayuda contextual. Manuales. Foros. Soporte técnico permanente.
- **Uso de estándares:** Adecuación a las especificaciones W3C. Transparencia respecto de la plataforma y navegador. Plugins universales. Compatibilidad con diferentes resoluciones de pantalla. Contenidos alternativos. Visualización de contenidos sin necesidad de descarga. Posibilidad de documentos PDF. Facilidad de impresión. Descarga de información en diferentes formatos.

Se han resumido aquí las principales heurísticas abordadas en algunas de las capas de MU_sa, de manera tal de ofrecer un detalle más acabado del tipo de análisis que se realiza en este modelo de evaluación propuesto. Sin embargo, debe aclararse que se deja de lado la explicación tanto de las sub-heurísticas como de los instrumentos concretos de recogida de datos, propuestos en MU_sa, ya que esto demandaría una mayor extensión en la descripción de cada capa.

A continuación se exponen las métricas utilizadas y mencionadas en la descripción de las capas de MU_sa. En particular, la tercera y cuarta capa incorporan, además de las heurísticas, un conjunto de métricas de usabilidad para poder evaluar algunos aspectos relacionados con la realización de las tareas propuestas a los usuarios en los escenarios. Cuando se hace mención a métricas se hace referencia a una medida cuantitativa del grado en que un sistema, componente o proceso posee un atributo dado. Las métricas seleccionadas poseen directa relación con el estándar ISO 9241-11 [14, 16, 19], y posibilitan medir la usabilidad del EVEA sin necesidad de estar involucrados en su desarrollo. Ellas son:

- **Eficiencia:** Porcentaje de tareas completadas
- **Efectividad:** Tiempo total invertido por el usuario en las tareas completadas. Tiempo total ocupado en errores. Tiempo total de aprendizaje. Frecuencia de uso de la ayuda o documentación.
- **Satisfacción:** Veces que el usuario expresa frustración o enojo, utilidad del producto. Satisfacción con respecto a las funciones y características. Percepción de que la tecnología da soporte a las tareas según las necesidades del usuario.

4. Aplicación de MUsa en un caso de estudio

Con el fin de poner a prueba el modelo propuesto, se abordó un caso de estudio, donde se utilizó MUsa para evaluar el EVEA SIAT (Sistema Informático de Apoyo a la Teleformación), en el contexto institucional de la Universidad Nacional de Río Cuarto (UNRC). El entorno virtual SIAT viene siendo desarrollado por la UNRC desde el año 2001, es utilizado por todas las carreras a distancia de la universidad, y como apoyo a las asignaturas de carreras presenciales [1].

Para la evaluación del EVEA en cuestión se utilizaron las cuatro capas de MUsa. Para la puesta en marcha de la evaluación se tomó contacto con diferentes espacios de trabajo de la Universidad Nacional de Río Cuarto (UNRC), y la Facultad de Informática de la Universidad Nacional de La Plata (UNLP). A continuación se describen las particularidades en relación a la aplicación de cada capa y los principales resultados obtenidos.

4.1 Capa 1

La primera capa del modelo fue llevada adelante por los cuatro evaluadores previstos en el modelo, dos de los cuales se consideraron dentro de la categoría de especialistas en educación y TIC, y otros dos especialistas en tecnología. Uno de los especialistas en tecnología se encargó de evaluar la flexibilidad tecnológica y organizativa del EVEA, mientras que el restante hizo lo propio con los estándares. Los dos especialistas en educación y TIC tuvieron a su cargo la evaluación de la flexibilidad pedagógico/didáctica y organizativa. La flexibilidad organizativa, en un principio, no estaba considerada en el modelo como parte de la evaluación del experto en tecnología, pero se decidió incluirla para tener otra mirada disciplinar y porque el especialista poseía muy buenos conocimientos acerca de las potencialidades organizativas del EVEA a evaluar. Una cuestión a resaltar es que el tiempo de evaluación se extendió por más de tres meses, mucho más de lo previsto. Otro aspecto importante es que el trabajo del experto dedicado a evaluar el seguimiento de estándares demandó el triple de tiempo que los demás, dado que la evaluación de estándares es un proceso lento y tedioso. A partir de la aplicación de esta capa del modelo, se obtuvo información muy valiosa para la mejora de compatibilidad y cuestiones de organización de SIAT. Todos los resultados forman parte de un informe institucional.

4.2 Capa 2

En la segunda capa es posible seleccionar la cantidad de evaluadores acorde a factores tales como tiempo, disponibilidad y presupuesto. Por este motivo, el modelo propone dos alternativas, una sencilla donde interviene un solo evaluador, y otra donde pueden intervenir de 3 a 5 evaluadores. Para este caso de estudio, se optó por seleccionar tres evaluadores con el fin de lograr una evaluación que recoja opiniones de más de un experto. Los evaluadores seleccionados fueron: uno de la Facultad de Ciencias Exactas, Físico-Químicas y Naturales de la UNRC, con experiencia en el uso del EVEA a analizar, y dos a la Facultad de Informática de la UNLP, sin experiencia en dicho EVEA. A su vez, para el caso de contar con varios expertos el modelo propone dos alternativas, una simplificada que evalúa el EVEA a nivel de tareas, y otra más exhaustiva que se basa en 21 escenarios de uso. En este caso se seleccionó la versión exhaustiva dado que la adopción de roles es más real, y posibilita realizar una suerte de "simulación" siguiendo la secuencia de escenarios planteada. Al contar con varios evaluadores es posible distribuir todos los personajes entre los mismos. Un cambio respecto al modelo original estuvo en la asignación de roles dado que se asignaron dos roles a cada evaluador, en lugar de uno, de manera tal que pudieran cubrirse los seis roles definidos (1 docente responsable, 2 tutores, y 3 alumnos). Otra diferencia radicó en la experiencia de los evaluadores. Se seleccionaron tres docentes con conocimientos en EVEA y mucha experiencia en utilización de computadoras y servicios de Internet (E). En cambio, el modelo plantea seleccionar expertos con diferentes grados y conocimientos en las temáticas relacionadas. Esto se debió a que las personas elegidas como expertos garantizaban compromiso y objetividad en la tarea. Su categorización se basó en el análisis del currículum vitae de cada uno.

Entre los obstáculos o problemas más importantes que debieron afrontarse en la aplicación de esta capa podemos destacar los siguientes: los evaluadores seleccionados no respondieron cabalmente al perfil propuesto en el modelo; el modelo no propone una forma para determinar el nivel de experticia del evaluador; el tiempo de evaluación se triplicó; los expertos seleccionados estaban dispersos en diferentes lugares geográficos. Su coordinación fue muy dificultosa y no se pudo realizar una puesta en común; el análisis pormenorizado de los escenarios permitió detectar algunos errores en su secuenciación; resultó engorroso solicitar que se utilicen varios navegadores y conexiones a Internet para realizar todos los escenarios; los escenarios resultaron ser algo dependientes entre sí (lo que hace un rol en un

escenario influye en lo que debe hacer otro rol en el mismo). Todos estos resultados fueron considerados para realizar ajustes en MUsa. Al mismo tiempo, toda la información obtenida en esta capa permitió realizar consideraciones para la evolución del EVEA SIAT.

4.3 Capa 3

Los usuarios participantes fueron solicitados a la Dirección de EAD de la Facultad de Ciencias Económicas de la UNRC. Se propuso un grupo de seis docentes y nueve alumnos, de acuerdo a los perfiles de usuario novato, medio y avanzado. En el caso de los docentes, se solicitó que 2 fueran novatos, 2 medios, y 2 avanzados; mientras que en el caso de los alumnos, 3 que fueran novatos, 3 medios, y 3 avanzados. Se acordaron dos sesiones de trabajo en días y horarios diferentes, y se seleccionaron dos laboratorios de usabilidad con características diferentes, uno con computadoras nuevas y otro con equipamiento de 10 años de antigüedad. Si bien la mayoría manifestó su predisposición a participar, resultó escasa la asistencia de los usuarios a cada sesión y complicada la coordinación de las mismas. De los usuarios convocados, concurrieron en total dos docentes y dos alumnos, en lugar de 6 y 9 respectivamente. Cada sesión estuvo a cargo de un observador para llevar el registro de lo acontecido durante la evaluación. A pesar de la escasa concurrencia, durante las sesiones se pudo apreciar que es necesario contar con un observador por cada dos evaluadores, en lugar de dos en total. La información recogida en esta capa fue de sumo interés para la evaluación del EVEA en cuestión.

4.4 Capa 4

La cuarta capa involucra una evaluación bajo condiciones reales de uso, es decir, en un contexto establecido y con usuarios reales. Para ello, se seleccionó un conjunto de usuarios finales pertenecientes al primer año de las tres carreras con modalidad a distancia de la Facultad de Ciencias Económicas de la UNRC, que estaban utilizando el EVEA, y poseían un contexto similar al establecido para los escenarios. Cabe destacar que las materias de primer año son las mismas para cualquier carrera. Otra alternativa era una evaluación simulada, pero se optó por la real dado que se contaba con aulas virtuales que estaban en pleno funcionamiento, y permitían acceder a resultados basados en el contexto real. La evaluación se llevó a cabo durante el segundo cuatrimestre de 2010 en las asignaturas Análisis Matemático II, Principios de Economía I, e Historia Económica y Social. Cada asignatura disponía de un

aula virtual dividida en dos comisiones cada una. El test remoto fue realizado, tanto por docentes como por alumnos, en su propio lugar de estudio o trabajo, de manera tal que fue posible evaluar el contexto de uso. Esta capa exige una cantidad mínima de evaluadores. En el caso de una evaluación simulada, son 42 usuarios, 15 de los cuales son docentes y 27 son alumnos. Para el caso de una evaluación real, como la realizada, esta cifra se reduce a la mitad, es decir, 7 docentes y 13 alumnos. La cantidad de docentes fue inferior a la exigida (4), mientras que la cantidad de alumnos alcanzó a cubrir el mínimo (15). Una diferencia importante con lo propuesto por MUsa fue que no se realizó un proceso de selección de los usuarios para conocer sus características básicas, y determinar su nivel de experticia y su contexto general, sino que se suministraron los cuestionarios por correo a todos los usuarios que habían participado de dichas materias. También se entregaron personalmente a los alumnos que concurrieron a la Universidad con motivo de un encuentro presencial de consultas.

Luego, se destinaron 2 meses para recibir las respuestas, y finalmente, el coordinador realizó un resumen de las opiniones, tomando como base los dos cuestionarios realizados por cada usuario, las huellas que dejaron dentro de las aulas virtuales durante el proceso de realización de las tareas, el registro de log de cada una, y el cuestionario completado por él mismo. Los problemas que se presentaron en cuanto a la aplicación de esta capa de MUsa estuvieron relacionados con la imposibilidad de cubrir la cantidad mínima de usuarios docentes, el no abordar una selección de usuarios acorde a su nivel de experticia (novato, medio, avanzado) y contexto general, y el largo tiempo de espera para lograr un número aceptable de respuestas. También es destacable el aporte realizado por una docente evaluadora que propuso modificar la escala de valoración de algunos instrumentos para mejorar la evaluación de las sub-heurísticas relacionadas con visibilidad, intuitividad, y topografía natural. En cuanto a la evaluación del EVEA en sí mismo, si bien los resultados han sido positivos en general, ha habido muchas sugerencias de los usuarios participantes que resultan de impacto para la evolución de este sistema.

5. Conclusiones y trabajos futuros

El modelo que se acaba de presentar está orientado a una evaluación de producto, es decir, sobre EVEAs que ya han sido desarrollados, y están en pleno funcionamiento. El modelo se denomina MUsa, dado que se trata de un Modelo centrado en la Usabilidad.

La evaluación se apoya en escenarios reales de uso, teniendo especial consideración por los alumnos y docentes, los objetivos que se proponen, las tareas específicas que realizan dentro del entorno durante las actividades de enseñanza y aprendizaje, el equipamiento e infraestructura que disponen, el lugar físico donde habitualmente se desenvuelven, y el entorno social en el cual están insertos. También puede facilitar la evaluación de nuevas versiones, o la comparación entre ellas.

MUsa se puso en práctica utilizando el EVEA SIAT como caso de estudio, en el contexto institucional de la Universidad Nacional de Río Cuarto. El objetivo principal fue testear el modelo dentro de un contexto conocido y acotado, para rescatar información útil de la experiencia adquirida durante la aplicación y la opinión de los expertos, observadores, docentes, y alumnos. Al mismo tiempo, los resultados de la evaluación fueron considerados en forma institucional para la mejora del sistema en cuestión.

A partir del análisis de los resultados, realizado luego del caso de estudio, se pudo observar que MUsa se comportó acorde a las expectativas. Sin embargo, hay cuestiones por pulir y mejorar a los efectos de dotar al modelo de mayor capacidad de adaptación ante diferentes circunstancias. Algunas de los problemas que necesitan ser abordados están relacionados con los tiempos destinados a la evaluación que fueron mayores a los esperados, el orden de algunas tareas pertenecientes a los escenarios que sufrió modificaciones, los instrumentos que recibieron críticas en las escalas de valoración, y en las adaptaciones que debieron realizarse en las capas 3 y 4 por no contar con la cantidad y calidad de evaluadores necesarios. Los próximos pasos a seguir tienen que ver con la revisión de todo el modelo de evaluación, su mejora de acuerdo a la información recogida a partir del caso de estudio y su presentación formal para su efectiva utilización.

Otro tema importante está relacionado con poder determinar el grado de usabilidad del EVEA evaluado con MUsa. El grado de usabilidad es una medida empírica y relativa. Habitualmente se basa en opiniones de expertos, pruebas de laboratorio, observaciones de campo, y pruebas de usuarios. El resultado depende de las metas planteadas para un contexto de uso determinado, o de una comparación con otras versiones o productos similares. En muchas ocasiones se intenta medir este grado de usabilidad mediante escalas cualitativas o cuantitativas. En MUsa, cada capa debería tener un resultado, como por ejemplo, muy aceptable, aceptable, no aceptable. La evaluación desfavorable de una capa tiene consecuencias importantes sobre la evaluación total, ya que difícilmente en las capas posteriores se mejore el grado de usabilidad del EVEA. Sin embargo, por la

amplitud del tema se lo considera como un aspecto abierto a trabajar una vez mejorado el modelo.

Referencias

- [1] Thüer, S.; Ferreira Szpiniak, A. Entornos Virtuales de Aprendizaje: Diseño de experiencias de usuario para la web 2.0. Conferencia Internacional ICDE – UNQ 2011 “Educación a distancia, TIC y universidad: calidad, equidad y acceso a la educación superior”. Bernal, Buenos Aires. Argentina (2011)
- [2] Adell, J. Selección de un entorno virtual de enseñanza / aprendizaje de código fuente abierto para la Universitat Jaume I. Centre d'Educació i Noves Tecnologies de la UJI con la colaboración del Servei d'Informàtica y del Gabinet Tècnic del Rectorat (2004)
- [3] Marquès Graels, P. Criterios de calidad para los sistemas de teleformación (2001)
- [4] Bevan, N. Guidelines and Standards for Web Usability (2005)
- [5] ISO/IEC 9126B1 Software engineering - Product quality - Part 1: Quality model. International Standards Organization (2001)
- [6] Keinonen, T. Usabilidad de los productos interactivos (2003)
- [7] Landauer, T. The Trouble with Computers: Usefulness, Usability and Productivity. Massachusetts: Massachusetts Institute of Technology (1995)
- [8] Nielsen, J. Usabilidad. Diseño de sitios web. Prentice Hall PTR (2000)
- [9] Nielsen, J. Usability 101: Introduction to Usability. Alertbox (2003)
- [10] Norman, D. Psicología de los objetos cotidianos. Editorial Nerea. ISBN: 84-89569-18-5 (1998)
- [11] Redish, J.; Dumas, J. A Practical Guide to Usability Testing (1999)
- [12] Schneiderman, B. Designing the User Interface: Strategies for Effective Human-Computer Interaction. Third edition, Addison Wesley Iberoamericana (1998)
- [13] Ruiz Morilla, J.J. ISO 9126 vs. SQuaRE (2009)
- [14] Tavis, D. ISO 13407 is dead. Long live ISO 9241-210!. Userfocus (2011)
- [15] ISO 9241-151 Guidance on World Wide Web user interfaces (2008)

- [16] ISO 9241-210: Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems (2010)
- [17] ISO/IEC TR 91264 Software engineering -- Product quality -- Part 4: Quality in use metrics (2004)
- [18] ISO/IEC 25000 Software Engineering -- Software product Quality Requirements and Evaluation (SQuaRE) -- Guide to SQuaRE (2005)
- [19] Rhodes, J. Usability Metrics (2000)
- [20] Zurita, G.; Sánchez, J.; Nussbaum, M. Taller Internacional de Software Educativo – TISE '99. Universidad de Chile (1999)
- [21] Ferreira Szpiniak, A.; Sanz, C. Hacia un modelo de evaluación de entornos virtuales de enseñanza y aprendizaje. La importancia de la usabilidad. Anales del XIII CACIC. Corrientes y Resistencia. Argentina (2007)
- [22] De Benito, B.; Salinas, J. Situaciones didácticas en los entornos virtuales de enseñanza-aprendizaje (EVEA) en la enseñanza superior: elaboración de un instrumento de análisis. Congreso EDUTEC (2006)
- [23] Lewis, C.; Rieman, J. Task-centred user interface design QA Practical introduction (1993)
- [24] Wharton, C.; Rieman, J.; Lewis, C.; Polson, P. The cognitive walkthrough method: A practitioner's guide. In usability inspection methods. Nielsen, J. and Mack, R.L. (Eds.), New York: John Wiley & Sons, pp.105-141 (1994)
- [25] Alva Obeso, M. E. Metodología de Medición y Evaluación de la Usabilidad en Sitios Web Educativos (2005)
- [26] Nielsen, J.; Molich, R. Heuristic evaluation of user interfaces. Proceedings of the CHI'90, ACM Press, pp 249-256 (1990)
- [27] Nielsen, J. Finding usability problems through heuristic evaluation. Proceedings of the CHI'92, ACM Press, pp 373-380 (1992)
- [28] Mack, R.; Nielsen, J. Usability inspection methods. ACM SIGCHI Bulletin-June, 2833 (1993)
- [29] Manchón, E. Evaluación heurística (o por expertos) de la usabilidad (2003)
- [30] Catalán Vega, M. Metodologías de evaluación de interfaces gráficas de usuario (2000)
- [31] Ferreira Szpiniak, A.; Sanz, C. Propuesta de un Modelo de Evaluación de Entornos Virtuales de Enseñanza y Aprendizaje. Escenarios de uso. V Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET'10), El Calafate, Santa Cruz. Argentina (2010)
- [32] Woodward, B. Evaluation methods in usability Testing (1998)
- [33] Hom, J. The Usability Methods Toolbox (2003)
- [34] Ferreira Szpiniak, A.; Sanz, C. Un modelo de evaluación de entornos virtuales de enseñanza y aprendizaje basado en la usabilidad. Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología. Nro 4 pp 10-21 (2009)

Dirección de Contacto del Autor/es:

Ariel Ferreira Szpiniak

Brasil 910

Río Cuarto

Argentina

e-mail: aferreira@exa.unrc.edu.ar

Cecilia V. Sanz

Calle 50 y 115

La Plata

Argentina

e-mail: csanz@lidi.info.unlp.edu.ar

Ariel Ferreira Szpiniak

Lic en Ciencias de la Computación. Docente de la Facultad de Ciencias Exactas, Físico-Químicas y Naturales de la UNRC.

Cursa la Maestría en Tecnología Informática Aplicada a Educación de la Facultad de Informática de la UNLP.

Cecilia V. Sanz

Dra. En Ciencias de la Facultad de Ciencias Exactas de la UNLP. Licenciada en Informática. UNLP.

Profesor Adjunto. Facultad de Informática de la UNLP.

Director y Profesor de la Maestría en Tecnología Informática Aplicada en Educación. Fac. Informática. UNLP.
